MODEL-DRIVEN DEVELOPMENT FOR PDS4 SOFTWARE AND SERVICES. J. Steven Hughes[1], Daniel Crichton[2], Stirling Algermissen[3], Michael Cayanan[4], Ronald Joyner[5], Sean Hardman[6], and Jordan Padams[7], [1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA, steve.hughes@jpl.nasa.gov, [2]daniel.crichton@jpl.nasa.gov, [3]stirling.algermissen@jpl.nasa.gov, [4]michael.cayanan@jpl.nasa.gov, [5]ronald.joyner@jpl.nasa.gov, [6]sean.hardman@jpl.nasa.gov, [7]jordan.padams@jpl.nasa.gov.

Introduction: Software and services that access Planetary Data System (PDS) PDS4 data products need to parse product labels to retrieve, interpret, and process the referenced digital objects. Under PDS4 a driving principle is that the product label provide all of the information necessary for these functions to be performed accurately. However, significantly more information is available in the PDS4 Information Model (IM)[1], the controlling document used to define, create, and syntactically and semantically verify the product labels. This additional information in the IM is made available for use, by both software and services, to configure, promote resiliency, and improve interoperability.

Overview of the PDS4 Information Model: As part of its information architecture, the Planetary Data System (PDS) developed the PDS4 Information Model [1]. This model captures the knowledge about planetary science data at several levels of specificity and allows humans and machines to communicate about the data. As shown in Figure 1, this knowledge has been translated and written to system files in several formats including XML Schema and Schematron, JavaScript Object Notation (JSON), Resource Description Framework (RDF), comma-separated values (CSV) files, XML Metadata Interchange (XMI), and OWL Web Ontology/Description Logic (OWL/DL). The information in these files provide information requirements that augment the system's functional requirements.

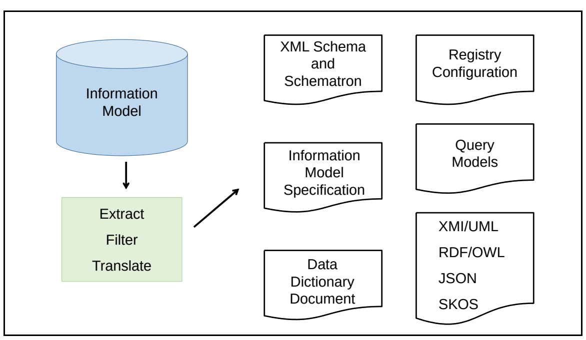

Figure 1 – Generated Artifacts

To address the challenges associated with change within the science community, a multi-level governance scheme was instituted. As Figure 2 shows, the IM has been partitioned into a single common, and several discipline and mission dictionaries.

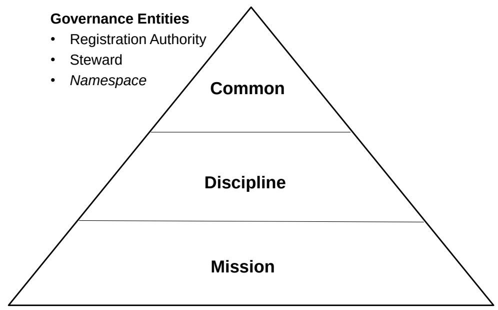

Figure 2 – Multi-level Governance Scheme

The partitioning of the IM reduces the impact of changes, to both the IM and the software and services that use the IM, by localizing the changes to the affected components. Since the common dictionary is designed to be stable, the relatively few changes to this dictionary are localized there. Changes to a discipline dictionary, for example cartography, again only impact that discipline. At the mission level the changes will be much more frequent but again remain localized to a single dictionary. Finally since the model is independent of the implementation, it is insulated from the relatively rapid rate of change in information technology. Software, developed to be configured by or that actively responds to the IM, benefits from this localization with the result that the system, software, and services are more resilient and insolated from changes to the IM. A change in the IM may not even require coding changes.

Multi-level governance also localizes interoperability. For example, the common dictionary enables interoperability across the entire community where each discipline and mission dictionary provides interoperability within a specific community.

Extending the Common Model: The PDS has been successful in developing a consistent and interoperable IM across its diverse disciplines and missions by first providing a stable common dictionary and then providing mechanisms to extend the IM in a controlled supervised enironment. The steward for each of the discipline and mission dictionaries create their individual models by encoding the model into an XML label. The XML label is validated for 'consistency' as part of ingestion into the existing IM (i.e., ensure conformity in the use of data types, units of measurement, and references across dictionaries). Since the initial release of the common model, several discipline and mission models have been released as shown in Figure 3.

Steward Name	Dictionary Description
Common	
Planetary Data System	PDS's common dictionary.
Discipline	
Cartography	Imaging Node's cartography dictionary.
Display	Imaging Node's display dictionary.
Geometry	Geometry dictionary.
Imaging	Imaging node's dictionary.
Ring-Moon Systems	Rings node's dictionary.
Spectral	Spectra dictionary.
Spectral Library	Spectral Node dictionary.
Planetary Plasma Interactions	PPI Node's Wave dictionary.
Mission	
BOPPS	BOPPS dictionary.
InSight	Insight dictionary.
LADEE	LADEE dictionary.
MGS	Mars Global Surveyor dictionary.
MVN	MAVEN dictionary.
OREX	OSIRIS-Rex dictionary.

Figure 3 – Model Dictionaries

Uses Cases: The following use cases illustrate how the exported system files are being used by PDS4 software and services.

Product Label Templates. The chosen implementation for PDS4 product labels is XML. The information model is converted to XML Schema files that contain the classes, attributes, constraints, and relationships defined for each type of data product in the information model. A data provider chooses the appropriate XML Schema files and using an XML editor generates a label template. The data provider then uses the XML editor to populate the template, producing a completed product label. These label templates are also used in product production pipelines to generate a series of product labels for sets of similar products.

PDS4 Validation. The PDS4 Validation tool uses the XML Schema files to validate PDS4 product labels during the ingestion process. This includes validating the order of elements in the label, whether a required element is present, and checking data types and minimum and maximum values. In addition, during generation of the XML Schema files, Schematron files are also written that contain rules to test constraints, for example checking that a value is a member of an enumerated list or that it has been formatted properly.

PLAID. The PDS Label Assistant for Interactive Design (PLAID) tool seeks to simplify and expedite the process of building a PDS4 label template with a simple step by step interface that does not require experience with XML, PDS4 Schemas and Schematron, or knowledge of the specific requirements of a PDS4 product label. PLAID is configured from a JSON formatted file that contains the contents of the PDS4 Information Model. Any change to the model is reflected in the tool.

Conclusion: The PDS4 Information Model, a set of information requirements for science data products, was developed by science experts in the planetary science community. These requirements specify the metadata required to sufficiently describe the data products so that they are scientifically usable by the Planetary Science community both now and into the future. The requirements are also machine readable and can be used to configure software and services that create, validate, and process the data products. Software and services that are written to be configured by or respond to the IM are more resilient, are more readily configurable, and result in reduced maintenance overall. Interoperability is established by the common and existing discipline and mission dictionaries and enhanced over time as new disciplines and mission dictionaries are designed and shared.

References:

[1] Hughes, J.S., Crichton, D., Hardman, S., Law, E., Joyner, R., Ramirez, P., PDS4: A Model-Driven Planetary Science Data Architecture for Long-Term Preservation, IEEE 30th International Conference on Data Engineering (ICDE), Chicago, IL USA, 2014.

Acknowledgements: This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.